



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Subfunctionalization of duplicated zebrafish pax6 genes by cis-regulatory divergence

Citation for published version:

Kleinjan, DA, Bancewicz, RM, Gautier, P, Dahm, R, Schonhaler, HB, Damante, G, Seawright, A, Hever, AM, Yeyati, PL, van Heyningen, V & Coutinho, P 2008, 'Subfunctionalization of duplicated zebrafish pax6 genes by cis-regulatory divergence', *PLoS Genetics*, vol. 4, no. 2, e29.
<https://doi.org/10.1371/journal.pgen.0040029>

Digital Object Identifier (DOI):

[10.1371/journal.pgen.0040029](https://doi.org/10.1371/journal.pgen.0040029)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

PLoS Genetics

Publisher Rights Statement:

Copyright: © 2008 Kleinjan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Subfunctionalization of Duplicated Zebrafish *pax6* Genes by *cis*-Regulatory Divergence

Dirk A. Kleinjan¹✉, Ruth M. Bancewicz¹✉^a, Philippe Gautier¹, Ralf Dahm²✉^b, Helia B. Schonthaler²✉^c, Giuseppe Damante³, Anne Seawright¹, Ann M. Hever¹, Patricia L. Yeyati¹, Veronica van Heyningen^{1*}, Pedro Coutinho¹

1 Medical Research Council (MRC) Human Genetics Unit, Western General Hospital, Edinburgh, United Kingdom, **2** Department of Genetics, Max-Planck Institute for Developmental Biology, Tübingen, Germany, **3** Department of Science and Biomedical Technology, University of Udine, Udine, Italy

Gene duplication is a major driver of evolutionary divergence. In most vertebrates a single *PAX6* gene encodes a transcription factor required for eye, brain, olfactory system, and pancreas development. In zebrafish, following a postulated whole-genome duplication event in an ancestral teleost, duplicates *pax6a* and *pax6b* jointly fulfill these roles. Mapping of the homozygously viable eye mutant *sunrise* identified a homeodomain missense change in *pax6b*, leading to loss of target binding. The mild phenotype emphasizes role-sharing between the co-orthologues. Meticulous mapping of isolated BACs identified perturbed synteny relationships around the duplicates. This highlights the functional conservation of *pax6* downstream (3') control sequences, which in most vertebrates reside within the introns of a ubiquitously expressed neighbour gene, *ELP4*, whose *pax6a*-linked exons have been lost in zebrafish. Reporter transgenic studies in both mouse and zebrafish, combined with analysis of vertebrate sequence conservation, reveal loss and retention of specific *cis*-regulatory elements, correlating strongly with the diverged expression of co-orthologues, and providing clear evidence for evolution by subfunctionalization.

Citation: Kleinjan DA, Bancewicz RM, Gautier P, Dahm R, Schonthaler HB, et al. (2008) Subfunctionalization of duplicated zebrafish *pax6* genes by *cis*-regulatory divergence. PLoS Genet 4(2): e29. doi:10.1371/journal.pgen.0040029

Introduction

Complex spatiotemporal control is required to regulate the expression of major tissue-specific transcription factors and signaling molecules which fulfill multiple roles in developmental patterning [1,2]. Accumulating evidence suggests that evolutionarily conserved non-coding elements orchestrate this process [3–5]. In addition to transgenic reporter studies in multiple species [6–12], human disease-associated genomic rearrangements (reviewed in refs [2,13]) and mutations [8,14], as well as natural and targeted deletions in mice [15–17], have contributed to our current understanding of how these regulatory elements control gene expression. Alteration of *cis*-regulatory function has been proposed as an important mechanism for evolutionary divergence [18]. Gene duplication provides an opportunity for such diversification without major penalties for loss of function [19]. About 350 million years ago, whole genome duplication occurred in an ancestor of teleost fishes, followed by variable loss of duplicated segments, and this process has been suggested to account for extensive speciation [20–23], so that half of all vertebrate species are teleost fish. Mechanisms for retention of duplicated copies remain controversial. Gain of new functions (neofunctionalization) has been widely proposed, but recently subfunctionalization, a more neutral partitioning of ancestral functions between duplicates, has been suggested [20,24]. A number of key transcription factors represented only once in mammalian genomes, are present as duplicates in zebrafish (*Danio rerio*). Often different loci are duplicated in other teleosts, such as medaka (*Oryzias latipes*) and pufferfish (*Fugu rubripes*). Where duplicate (or triplicate) genes have been examined, they reveal functional divergence,

although the mechanisms have generally not been explored [25].

Our interest in developmental eye anomalies in humans and model systems, led us to map the homozygous viable and fertile zebrafish microphthalmia mutation *sunrise* (*sri*) [26]. The original mutant was isolated in the Tübingen ENU mutagenesis screen [27]. We identified a deleterious missense mutation in the homeodomain of the *pax6b* gene. To assess gene function in the context of observed phenotype, this

Editor: Wayne N. Frankel, The Jackson Laboratory, United States of America

Received: June 27, 2006; **Accepted:** December 21, 2007; **Published:** February 15, 2008

Copyright: © 2008 Kleinjan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abbreviations: dpf, days post-fertilisation; E60, highly conserved enhancer 60 kb downstream of *PAX6* promoter; *ELP4/Elp4/elp4*, human, mouse and zebrafish elongation protein 4, regulatory component of RNA polymerase 2 complex; ENU, ethyl-nitroso-urea (alkylating mutagen); hpf, hours post-fertilisation; LG, linkage group; *MHB*, mindbrain-hindbrain boundary; *P/EE*, pancreas/ectodermal enhancer element; *PAX6/Pax6/pax6*, human, mouse and zebrafish paired domain and homeodomain protein 6 gene; *RCN1/Rcn1/rcn1* human, mouse and zebrafish reticulocalbin 1; *sri*, sunrise small eyed phenotype in zebrafish; *WT1/Wt1/wt1* human, mouse and zebrafish Wilms tumour predisposition gene

* To whom correspondence should be addressed. E-mail: v.vanheyningen@hgu.mrc.ac.uk

✉ These authors contributed equally to this work.

^a Current address: The Faraday Institute for Science and Religion, St Edmund's College, Cambridge, UK

^b Current address: Center for Brain Research, Medical University of Vienna, Vienna, Austria

^c Current address: Research Institute of Molecular Pathology, Vienna, Austria

Author Summary

Studying the zebrafish small eyed mutant “sunrise,” we identified the causative amino acid change in the *pax6b* gene. This mutation leads to reduced DNA binding capacity. There are two closely related *pax6* genes in zebrafish, *pax6a* and *pax6b*, which arose following a whole-genome duplication event about 350 million years ago; they map to different chromosomes. Each copy is now associated with a different subset of the neighbouring genes found associated with all vertebrate single-copy *Pax6* genes. The expression patterns of *pax6a* and *pax6b* have diverged from each other since the duplication event. Some division of labour has emerged: *pax6b* is less widely expressed in the brain than *pax6a*, but only *pax6b* is found in the developing pancreas. Multiple evolutionarily conserved regulatory elements (enhancers) control these expression patterns, which can be recapitulated in transgenic animals. Some enhancer elements lie more than 150 kb outside the transcribed gene region, inside the introns of unrelated neighbouring genes. Such juxtaposition imposes the need to conserve gene order in many vertebrate species. Genome duplication releases the constraint for retaining all neighbouring genes. Thus, *pax6a* has lost the coding region of its immediate neighbours, although it retains most of the brain-specific regulatory domains. Duplication also allows some orderly changes in the exact role of each regulatory component, as long as the two duplicates can, together, ensure the complex expression pattern required for complete function. We demonstrate functional loss of a brain element downstream of *pax6b*, while an upstream pancreas enhancer element has evolved in a more complex way.

finding triggered a comparison of expression patterns and regulatory control elements for the duplicated *pax6a* and *pax6b* genes, with each other and with other vertebrates where generally only a single orthologue is found.

PAX6 is a highly conserved protein, with paired- and homeodomain DNA-binding regions; it functions as a transcription factor with a major role in eye and brain development from *Drosophila* to humans [28,29]. Homozygous loss-of-function mutations lead to absence of the eye and lethal brain anomalies in humans, mice and flies [30–33]. Heterozygous null mutations (haploinsufficiency) are associated with human aniridia and murine microphthalmia (Small eye); missense changes lead to related, sometimes more severe, eye defects [29]. Associated brain anomalies have also been described in both human and mouse heterozygotes [34,35]. Its expression pattern suggested, and whole animal functional studies have shown, that Pax6 also plays an essential role in the development and adult maintenance of pancreatic endocrine cells [36]. When the conditionally targeted *Pax6* gene is homozygously inactivated in the mouse pancreas at E9.5, the pups die of severe diabetes between postnatal days 3 and 6.

The complex spatiotemporal expression of Pax6 is controlled by an extensive downstream regulatory region [6] whose existence was heralded by haploinsufficient aniridia-associated chromosomal disruptions positioned more than 150 kb downstream of *PAX6* [37]. DNaseI hypersensitivity analysis and strong, consistently patterned enhancer function in reporter transgenic animals confirmed and elaborated the predicted regulatory element organization [6,7,38]. In many cases regulatory element function can be predicted by non-coding region genomic sequence conservation in mammals or, more broadly, in vertebrates [7,10,12,38–40].

It was against this background that we examined the *sri* mutant and the regulation of the two *pax6* co-orthologues, to gain insight into the evolutionary conservation and divergence of these duplicate genes.

Results/Discussion

Mapping and Identification of *pax6b* as the *sri* Gene; Functional Analysis of Wild Type and Mutant Proteins and Definition of the Phenotype

Detailed analysis of eyes in *sri* mutant zebrafish [27], revealed a variable, but fully penetrant recessive phenotype, with abnormal lens and corneal structure leading to reduced eye size (Figure 1A). Affected homozygotes are viable and fertile, and this allowed outcrossing to the wild-type line WIK for mapping [41,42]. The uniformly heterozygous F1 siblings from this cross were interbred and the recessive phenotype was recovered, permitting the collection of phenotypically classified individuals for mapping. Standard techniques [41,42] were used to map the *sri* locus to a region, between 29.7 and 33.2 cM from the top of linkage group 7 (LG7) (Figure S1A). *pax6b* [43] was identified as a strong candidate gene. Sequencing of *sri* homozygotes and heterozygotes identified a leucine to proline missense mutation in a highly conserved residue of the *pax6b* homeodomain (Figures 1B, S1B, and S1C). The L244P alteration, in the first helix of the homeobox, is predicted to affect protein structure and function significantly [44]. A luciferase reporter assay was used to assess the regulatory capabilities of the full length Pax6a and Pax6b wild type proteins, and that of the L244P Pax6b mutant. Activation by the two wild type co-orthologues through the P3 homeodomain target [45] (Figure 1D), or the CD19 paired domain target [46] (Figure 1E), was not significantly different, in contrast to the original observations by Nornes et al [43]. However, the mutation abolished activation through the P3 homeodomain target and severely compromised CD19 paired domain function (Figure 1D and 1E) [47,48]. Binding of mutant protein to both high-affinity homeodomain target sequences P2 and P3, with different spacing between the twin sites [45], is also reduced in an electrophoretic mobility shift assay (Figure 1F). Following the identification of the mutation, the frequencies of the three expected genotypes were assessed after crossing obligate heterozygotes and found not to deviate from the expected 1:2:1 Mendelian ratios at 5 dpf, confirming the full fitness of the homozygotes at this stage (Figures 1C and S2). Survival into adulthood is good; *sri* fish are maintained as homozygotes. The viability, fertility and relatively mild eye phenotype of the homozygous *pax6b* mutant fish, led us to investigate further the role of this gene, and its co-orthologue *pax6a*, which maps to LG25 (http://zfin.org/cgi-bin/mapper__select.cgi) and Supporting Information).

Comparison of the *pax6a* and *pax6b* Expression Patterns and Analysis of the *sri* Mutant

Overlapping divergent expression patterns have been reported for *pax6a* and *pax6b* (previously named *pax6.1* and *pax6.2* respectively) [43,49]. *pax6b* is the “minor” co-orthologue, expressed in the developing eye (retina and early lens placode), and also in thin strips of dorsal diencephalon, at the midbrain-hindbrain boundary (MHB), and in the pancreas. *pax6a* is expressed in the lens and retina, as well as more

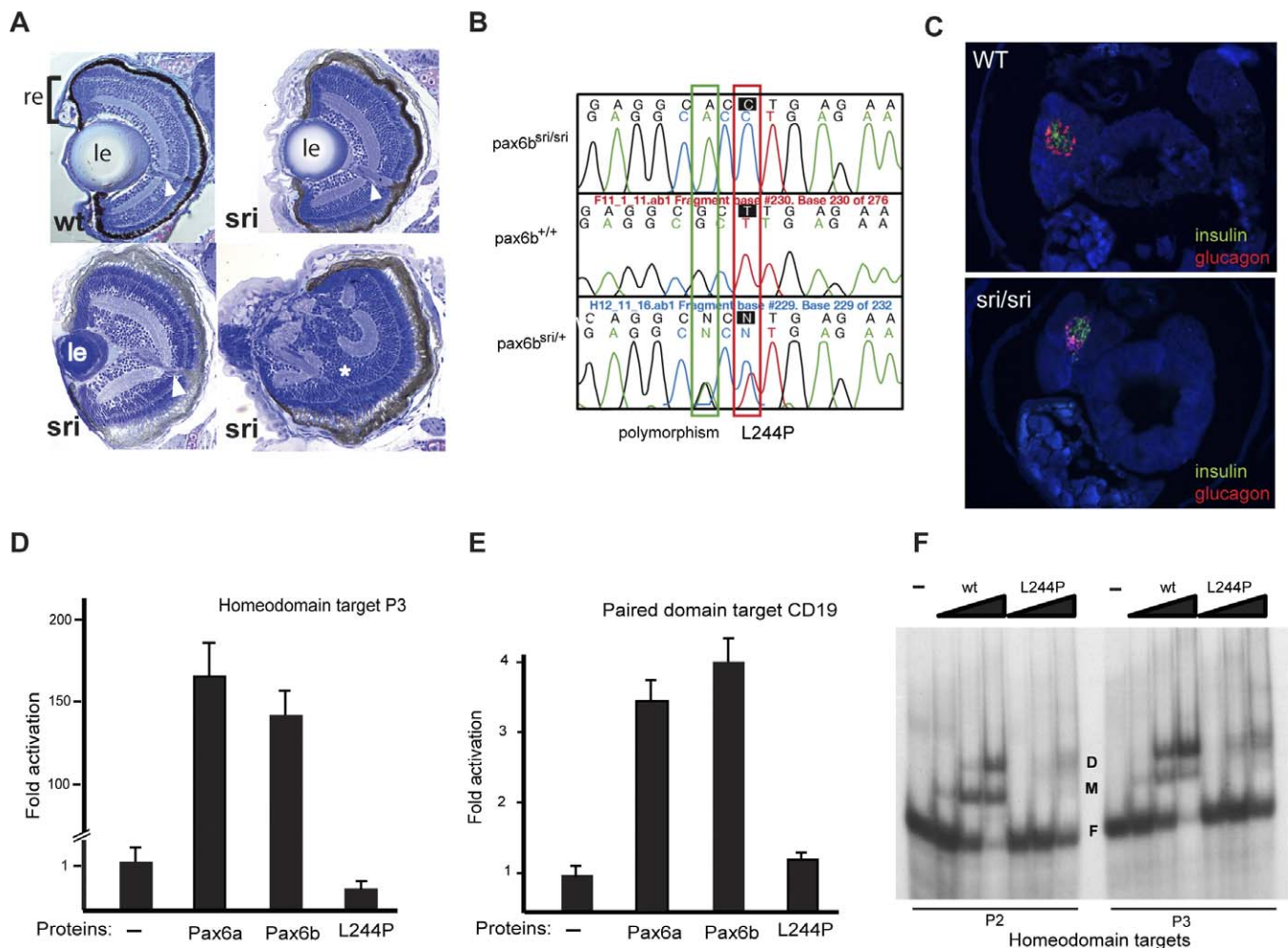


Figure 1. Characterisation of the *sunrise (sri)* Mutant

(A) Sections through wild type and homozygous *sri* larval eyes at 5 dpf, showing variable severity from mildly reduced lens size to almost complete absence of lens and retinal malformation (*). le, lens; re, neural retina and retinal pigment epithelium (black layer); arrowhead, optic nerve. (B) Sequence traces from *pax6b* homeodomain revealing the T to C mutation resulting in the L244P mutation. The G to A third position change two bases earlier is a polymorphism between the WIK and Tü strains. (C) Expression of known *pax6* targets, glucagon and insulin, in pancreas of wild type (wt) 5 dpf fish. Expression is maintained in *sri* despite the loss-of-function homeodomain missense mutation in *Pax6b*, the sole *pax6* gene expressed in the pancreas. (D–F) Functional analysis of the *Pax6b^{sri}* (L244P) protein in comparison with wild type *Pax6b* and *Pax6a*, using luciferase reporter assays in HeLa cells and EMSA. (D) *Pax6a* and *Pax6b* proteins drive luciferase expression at comparable levels under the control of the P3 homeodomain promoter; the L244P mutant of *Pax6b* protein fails to activate P3. (E) *Pax6a* and *Pax6b* proteins drive comparable luciferase expression levels through the CD19 paired domain target promoter; the L244P mutant of *Pax6b* protein has significantly reduced activity. (F) Gel-retardation assay to analyse the binding capacity to homeodomain target binding sites P2 and P3 demonstrates reduced affinity of the L244P mutant protein compared to wild type (wt) *Pax6b*. D, dimeric protein-bound; M, monomeric protein-bound; and F, free labeled target oligonucleotide. Protein concentrations used 0.5, 1.5, and 4.5 μM (lanes from left to right) with 5 nM oligonucleotides. doi:10.1371/journal.pgen.0040029.g001

widely in the developing telencephalon, diencephalon, hind brain, and spinal cord, although not in developing pancreas [49]. The expression pattern of the co-orthologues was confirmed by RNA in situ analysis at 24 and 32 hpf (Figure 2A–2J). There is significant overlap between the territories of the co-orthologues, clearly illustrated in the sections showing *pax6a* and *pax6b* expression (Figure 2I and 2J).

We set out to verify that the L244P mutation is causative for the *sunrise* phenotype. First, *pax6b* morpholino injections, using two different morpholinos, resulted in reduced eye size, a phenotype overlapping the *sri* phenotype, but somewhat more severe, as total eye size reduction (Figure 3A and 3B)

rather than just small lens size [27] (Figure 1A) was observed. Morpholino injection was found to reduce protein expression proportionately to dosage (Western blot data not shown), ultimately leading to a null phenotype, while the homeodomain missense mutation may possess some residual function, even if DNA-binding capacity was strongly diminished by the mutation. This differential effect of missense and null mutations is also illustrated by the distinct phenotypes observed when patients with heterozygous missense mutations in the paired domain are compared with haploinsufficient cases ([29,50,51] and van Heyningen et al unpublished). In addition to the reduced eye phenotype, developmental

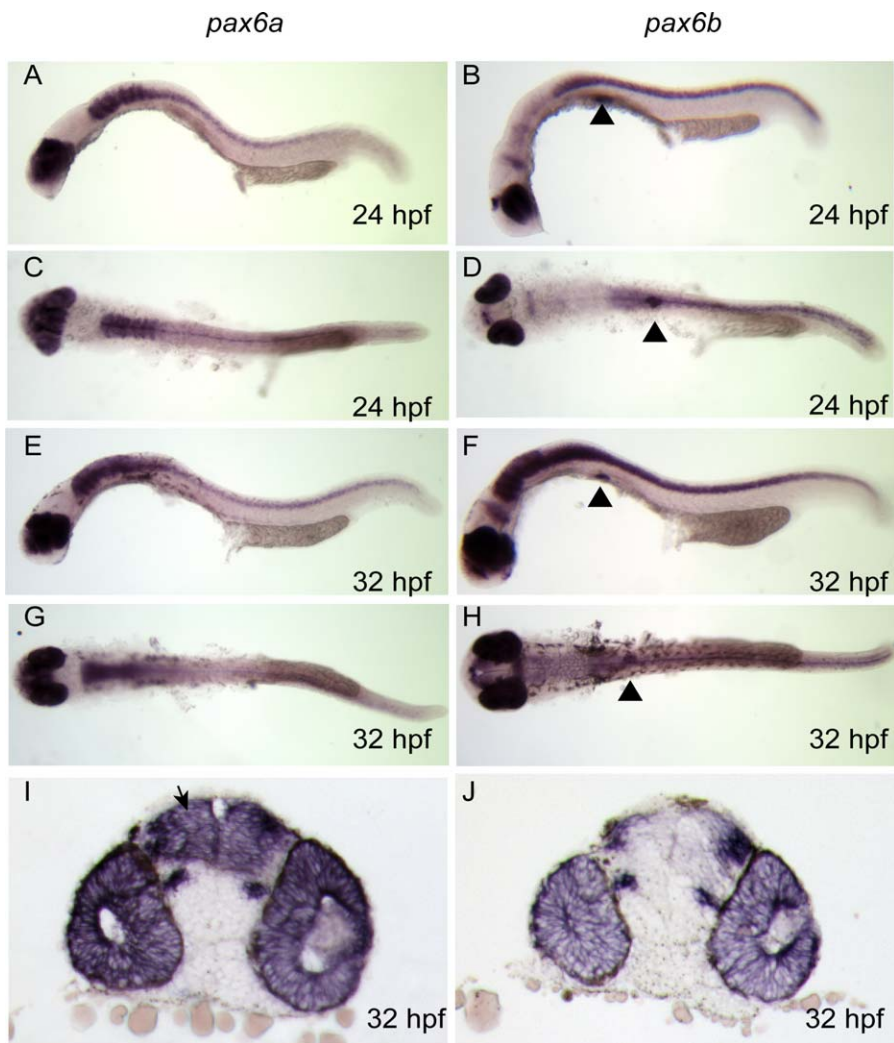


Figure 2. *pax6a* and *pax6b* Expression Analysis by Wholemount RNA *in situ* Hybridisation during Early Zebrafish Development

pax6a expression is shown in (A) and (C) at 24 hpf; (A) lateral view showing expression in eye, telencephalon, diencephalon, hind brain, and neural tube (C) dorsal view; (E) and (G) lateral and dorsal view at 32 hpf, showing continuing pattern of expression, (I) Vibratome coronal section of wholemount stained embryo expressing *pax6a* in telencephalon (arrowed), diencephalon, retina and lens. (B) *pax6b* expression is seen in lateral view showing eye and optic tectum at 24 hpf, pancreas expression is highlighted using black triangles; (D) dorsal view at 24 hpf. (F) and (H) lateral and dorsal. Increasing hindbrain and neural tube expression of *pax6b* is revealed by 32 hpf embryos. (J) predominantly eye expression of *pax6b* is seen in coronal section at 32 hpf.

doi:10.1371/journal.pgen.0040029.g002

delay is observed if the fish are followed for 48 hpf (Figure 3B), although expression of *pax6a* mRNA continues (Figure 3C and 3D). Simultaneous injection of *pax6a* and *pax6b* morpholinos disrupts eye development leading to microphthalmia and general developmental delay (data not shown).

Rescue of the *sri* phenotype was attempted, using capped mRNA injection into 1–2 cell stage embryos [52,53], but this led to severe developmental delay at a later stage, with reduced eye size, at mRNA levels normally used for such studies (Figure 3E). We postulate that this is because the requirement for functional Pax6 protein is both tissue-specific and highly dosage sensitive, and neither of these aspects can be controlled during mRNA-rescue by injection. Protein carrying a single altered amino-acid is usually expressed normally, probably with some intact functional domains. It is likely to be present at regular or even at elevated levels, as protein with reduced function may not be capable of driving the negative autoregulation that has been

observed for Pax6 in mice [54,55]. Titration of mRNAs in wild type one- or two-cell embryos was repeated three times and the data combined (Table 1). This revealed that abnormal phenotypes begin to be elicited even with 10 pg injections of wild type *pax6b*, and increasing the amount of mRNA increases the severity score (Table 1). However the mutant L244P Pax6b-encoding mRNA consistently produces less severe phenotypes than the wild type at equivalent levels (Figure 3E; Table 1). Thus, although we were unable to rescue the effect of the missense mutation using capped mRNA injection into 1–2 cell stage mutant embryos, we show that increased dosage of wild type *pax6b* leads to a deleterious phenotype, and that the over-expression phenotype is less severe with an equal amount of mRNA encoding the L244P mutant protein, thus providing additional indirect evidence that this partial loss-of-function mutation is the underlying cause of the *sri* phenotype.

In the mouse *Pax6* expression is observed in fetal and adult

Table 1. Injection of Capped mRNA into Wild Type (Tupfel Long Fin) Fish

Wild Type Recipients				Eye Defects, Normal Body Size				Eye Defects, Dwarf Body			
<i>pax6b</i> mRNA injected (pg)	Dead	Number Alive	Normal Eyes	Two Eyes (Reduced)	One Eye	No Eyes	Percentage Eye Defect	Two Eyes	One Eye	No Eyes	Percentage Eye and Body Defect
Wild type (400)	24 (69%)	11	9 (25%)	0	0	0	0%	0	1	1	6%
L244P (400)	12 (30%)	28	12 (30%)	11	0	0	27%	4	0	0	10%
Wild type (25)	4 (12%)	28	8 (25%)	15	1	0	50%	0	1	2	9%
L244P (25)	6 (15%)	35	6 (15%)	9	2	2	32%	4	3	9	39%
Wild type (10)	7 (13%)	47	22 (40%)	13	5	3	38%	1	2	1	7%
L244P (10)	6 (11%)	51	48 (84%)	3	0	0	5%	0	0	0	0%

doi:10.1371/journal.pgen.0040029.t001

pancreas [36]. In zebrafish only *pax6b* is expressed in the developing pancreas, at least up to 48h (Figure 2), as previously reported [49]. We have also explored *pax6* expression in adult zebrafish and demonstrate by RT-PCR that only *pax6b* is expressed in the pancreas isolated from 6 month old wild type and *sri/sri* fish, while eyes from the same individuals express both *pax6a* and *pax6b* (Figure S3). We also showed that *pax6a* expression is not induced in the pancreas of the sunrise mutant. However, if only *pax6b* is expressed in the zebrafish pancreas, then the continuing normal expression of insulin and glucagon in the single-islet zebrafish pancreas of *sri* homozygotes (Figure 1C), is surprising. All known homozygous null and most missense mutations of *Pax6* in mice [56–58], as well as most reported compound heterozygotes in humans [31], and van Heyningen et al, unpublished data) die immediately postnatally, if allowed to come to term. However, there is evidence, in both mouse and humans, that rare homeodomain missense mutations in *Pax6* produce significantly milder phenotypes than nonsense mutations or paired-domain missense changes [29,51,56,59,60]. The significant but not complete abolition of homeodomain target DNA-binding, seen when the L244P mutation is assessed using EMSA (Figure 1F), is in line with this theory. Another possibility is that the paired domain, and not the homeodomain, fulfils the critical role in driving gene expression in the pancreas [61], where insulin, glucagon and somatostatin are all regulated by the –5a form of the *Pax6* paired domain [62]. Although the luciferase reporter studies reveal that the *sri* homeodomain mutation severely limits paired domain binding as well (Figure 1E), this may be a cell-line-specific effect [63]. However, the converse has also been reported, namely that paired domain mutations disrupt homeodomain binding [47,48], suggesting that the *Pax6* protein-DNA binding interaction requires collaboration between the paired- and homeo-domains. Interestingly, we recently identified (van Heyningen et al, unpublished) compound heterozygosity for two different *PAX6* mutations (a premature truncation and a missense change at a completely conserved arginine just N-terminal to the homeodomain) in a child who is developing normally apart from anophthalmia, suggesting that missense changes in the homeodomain region may be viable. The missense mutation carrying mother of this child has a severe eye phenotype. A similar surviving compound heterozygote case has also been reported [64].

Evolution of Synteny Relationships around *pax6a* and *pax6b*

To gain further insight into the evolution of the duplicated zebrafish *pax6a* and *pax6b* genes, we assessed their conservation and divergence through extensive sequence comparison over coding and non-coding domains, including multiple functionally defined *cis*-regulatory elements [6,7,9,65], and through regional synteny studies. Human reticulocalbin (RCN1) was mapped between PAX6 and WT1 (Wilms tumour predisposition gene), using somatic cell hybrids [66]. Synteny conservation with mouse was noted [66], and later extended to Fugu [67]. Subsequently, the transcriptional elongation protein 4 (ELP4), a component of the elongating RNA polymerase II complex, was identified as the neighbour gene downstream of PAX6 [68]. In addition to well-defined regulatory elements upstream of promoters P0 and P1 [39, 65,69,] and within introns [9,65,70,71], a number of conserved elements required for correct spatiotemporal expression of PAX6 were found to map downstream of PAX6 within introns of ELP4, thus creating a requirement for obligate synteny conservation between the two genes [67,68]. Several of the conserved elements have been shown to function as transgenic enhancers capable of driving lacZ reporter expression, in spatiotemporal patterns corresponding to subsets of the complete *Pax6* expression pattern [6,7,38].

The suggestion that the two copies of zebrafish *pax6* arose through whole genome duplication [72] led us to explore the current synteny relationship of each co-orthologue, with the expectation that we would find two genomic copies of the known closely linked human markers, *WT1*, *RCN1* and *ELP4* (Figure 4A). However, only one copy has been identified for each of these three closely linked genes in zebrafish at one or other of the expected syntenic sites [73] on either LG7 or LG25 (Figure 4A). Genomic localisation was confirmed for all three genes by mapping on the T51 radiation hybrid map [41,42], using specific primers designed from available zebrafish sequence data in public genome annotations. Assignments were confirmed on multiple BAC clones we had isolated for each *pax6* orthologue, including DKEYP-46C10 and DKEY-157G7 containing *pax6a* and *pax6b* respectively (Figures S4 and S5). These BACs were part of the zebrafish genomic sequence assembly, and are integrated into regional maps for LG7 and LG25, verifying the syntenic relationships on these linkage groups.

Zebrafish *wt1a* was linked to *z14229* as the nearest marker, upstream of *pax6a* on LG25 (Figures 4A and S4), showing well-

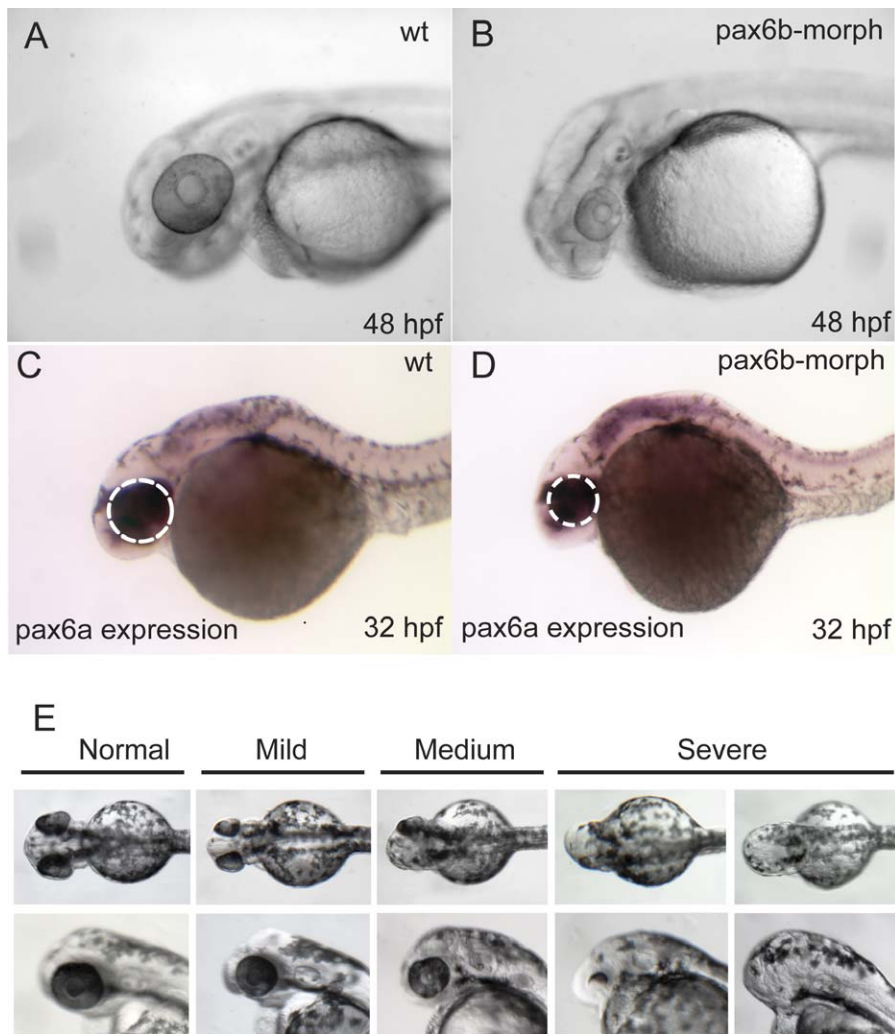


Figure 3. Analysis of *pax6b* Function in Zebrafish, with the Aim of Validating the Effect of the *pax6b* L244P Missense Mutation in *sri* Fish
(A–D) Phenotype elicited using *pax6b* morpholinos: (A) Uninjected wild type fish at 48 hpf; fish injected with the same amount of a control scrambled morpholino also produced no effect.
(B) Substantially reduced eye size and generally delayed development are observed by 48 hpf in response to injection of either of the two *pax6b* morpholinos.
(C and D) At 32 hpf reduction in eye size is observed (white dotted outlines) in otherwise normal looking embryos.
(E) The effects of injecting capped mRNA at levels that generally rescue the phenotype illustrating the spectrum of phenotypes to bilateral anophthalmia, observed when *pax6b* capped mRNA is injected at levels generally used for phenotype rescue experiments in zebrafish.
doi:10.1371/journal.pgen.0040029.g003

conserved exon structure and sequence, in agreement with recently reported data [74]. A related *wt1* sequence was also found on LG18, close to marker *zc50f22*, but with no neighbouring *pax6* sequence (data not shown), although this region of LG18 is known to harbour a human chromosome 11 synteny region [73]. Sequences corresponding to the coding exons of *rcn1* and *elp4* were only found flanking *pax6b* on LG7 (Figures 4A and S5). This pattern of partial synteny conservation is best illustrated with a PIP plot [75] using the compact *Fugu* sequence as the baseline (Figure S6), which reveals the exonic sequence conservations very clearly: showing that *wt1a* is linked to *pax6a*; and *rcn1* and *elp4* to *pax6b* [67]. A similar PIP plot with the chick sequence as the baseline (Figure S7) shows unequivocally that the conserved elements associated with the downstream region of human and mouse *PAX6*, mapping within a number of the *ELP4* introns, are still present downstream of zebrafish *pax6a*,

despite the loss of *elp4* exonic conservation. In contrast, most downstream elements have been lost from the *pax6b* locus where *elp4* has been maintained. This suggests that the whole syntenic region was ancestrally duplicated, but, in the absence of strong selective constraints, the *elp4* duplicate copy associated with *pax6a* was lost, presumably through multiple mutations (non-functionalisation), while the proposed *pax6b* regulatory elements, formerly residing in its introns, have been selectively maintained. These findings clearly imply that these conserved elements are not involved in *elp4* RNA processing, a possible role suggested for intronic conserved sequences [76].

Evolutionary Divergence of Regulatory Elements Associated with the Co-orthologues—Bioinformatic and Transgenic Reporter Analysis in Mouse and Zebrafish

Focusing on the evolving role of the duplicated zebrafish *pax6* genes, we observed multiple examples of subfunction-

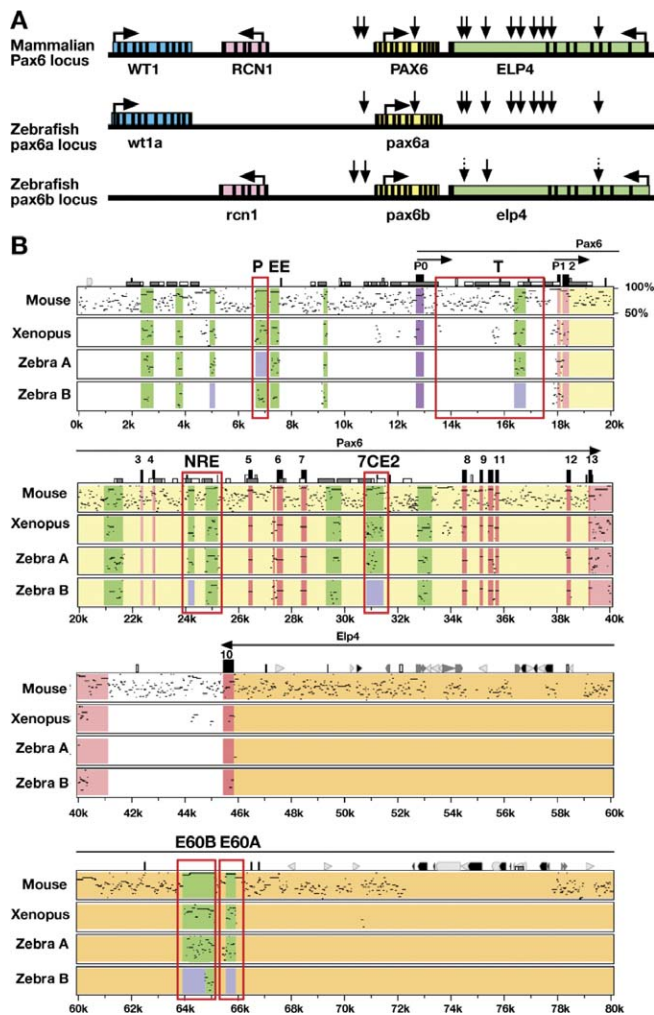


Figure 4. Comparative Linkage and Sequence Analysis of *pax6* Loci (A) Syntenic relationships of *pax6* with its nearest neighbours. In mammals *pax6* is flanked by *rcn1* and *elp4* genes with *wt1* located further upstream. Horizontal arrows show direction of transcription. Vertical arrows indicate the positions of a subset of the *pax6* cis-regulatory elements, many located downstream of *pax6*, within introns of the adjacent *elp4* gene. Dotted arrows denote partially conserved elements. Zebrafish LG25 retains *pax6a* synteny with *wt1a* and with the majority of *pax6* cis-regulatory elements, but *rcn1* and *elp4* coding exons have been lost. On LG7 the *pax6b* locus and *rcn1* and *elp4* exons have been retained, but many of the *pax6b* control elements, as well as the *wt1* homolog, have been lost. (B) PIP plot comparing the *PAX6* locus in human with mouse, *Xenopus* and zebrafish *pax6a* and *pax6b*. Exons are shown in red, untranslated regions in pink, conserved cis-regulatory elements are shown in green, while absence of significant sequence conservation is indicated by a light blue box, and promoter P0 is demarcated in purple. The cis-regulatory elements discussed in the text are surrounded by open red rectangles. Enhancers: P, pancreas; T, telencephalon; NRE, neuroretina; 7CE2, diencephalon; E60B, ultraconserved enhancer with multiple brain specific functions (D. McBride, D. A. Kleinjan, and V. van Heyningen unpublished data); E60A, optic cup, early telencephalon and diencephalon P3 region. doi:10.1371/journal.pgen.0040029.g004

alization, or job-sharing, between *pax6a* and *pax6b*. The differential expression patterns of these co-orthologues are reflected by loss and retention of relevant conserved regulatory elements (Figure 4B).

In the *PAX6* downstream region, in the large intron 9 of *ELP4*, we identified a new regulatory element, E60, showing 94% nucleotide sequence identity between human and

mouse over 1400 bp, through sequence comparisons by PIP plot (Figures 4B and S7). Two distinct subcomponents, E60A and E60B, can be distinguished, as shown (Figures 4B and 5A). E60B contains one of the largest ultra-conserved elements: 601bp of complete human-mouse identity [77]. Detailed VISTA analysis [78] revealed clear conservation of E60B in *Xenopus* and to some extent in both zebrafish *pax6a* and *pax6b* (Figure 5A). For E60A only *pax6a* showed discernible conservation (Figures 5A and S8A), while *pax6b* (Figure S8B) has no homologous conserved sequence. There have been recent suggestions that enhancer activity may be maintained even where no sequence conservation can be observed between mammals and teleosts [79]. We therefore set out to assess the function of the E60A region in parallel studies in both mice and zebrafish. A set of transient and permanent mouse transgenic reporter lines was produced, with the human E60A region (Figure S8A) as the enhancer, using previously described methods [6,7,9] (Figure 5B–5E). The transgenic mice show lacZ expression consistently in the early retina, in the telencephalon, and in the future ventral thalamic region of the diencephalon (Figure 5B), with minor ectopic expression seen only occasionally, for example, in the limb (Figure 5D and 5E). It is not clear whether the neural tube staining is due to the Hsp70 minimal promoter used.

The zebrafish E60A region (Figure S8A and S8B) from *pax6a* or *pax6b* (located using the conservation at the linked E60B element as an anchor point) was inserted, using Gateway cloning, into a Tol2 transposon based vector containing a cFos minimal promoter-GFP reporter cassette (Figure 5A) for co-injection with Tol2 mRNA, as described [53] (Figure S8C). The E60A region from *pax6a* gives rise to transgenic fish expressing GFP in the optic vesicle and the telencephalon (24/38 surviving transgenic embryos showed this expression pattern at 32 hpf in a typical experiment) (Figure 5F and 5H). No consistent or significant expression was seen in any of 23 surviving transgenic embryos similarly injected with the homologous *pax6b* construct (Figure 5G and 5I). In this instance, therefore, maintenance or loss of sequence conservation correlates strongly with the presence or absence of tissue-specific enhancer function across multiple species.

Next, using zebrafish reporter transgenics, we assessed the roles of the compound element, P/EE, situated upstream of the two *Pax6* promoters, P1 and P0 in the mouse, and previously shown to include an ectodermal enhancer (EE) involved in regulation of *Pax6* expression in lens, cornea and lachrymal gland development, and a pancreatic enhancer P (Figures 4B and S9) [36,65,80,81]. Conditional knockouts had confirmed the role of P/EE in both lens and pancreas development [36,80,81]. PIP plots [75] illustrate the conservation of the 5' pancreas component [65] in the upstream region of *pax6b* and its absence in *pax6a* (Figures 2, 4B, and S7), which is not expressed in pancreas (Figure S3) [43,49]. In contrast EE and an additional neighbouring element appear to be conserved in association with both *pax6a* and *pax6b*. The P/EE element from *pax6a* and *pax6b* failed to elicit GFP expression when either was inserted into the same c-fos minimal promoter containing vector as the two zebrafish E60A elements (Figure 5A). We therefore engineered a new Tol2 based vector system containing a two-way Gateway destination cassette which can accommodate a reporter

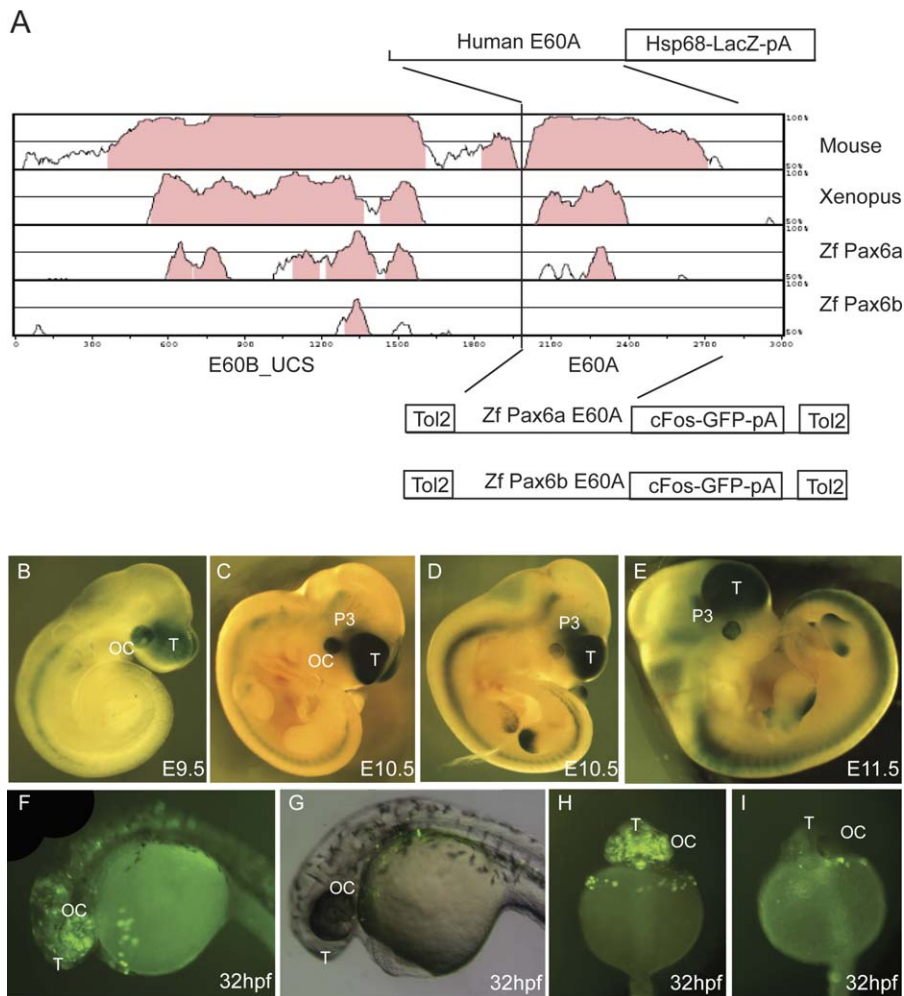


Figure 5. Enhancer Capacity of the E60A Conserved Element Assessed in Transgenic Mice and Zebrafish.

(A) Vista plot of the region around the highly conserved region E60+, located 47 kb downstream from the PAX6 P1 promoter. Human sequence is used as the base for pairwise comparison with *pax6* loci from mouse, *Xenopus tropicalis*, and zebrafish (*pax6a* and *pax6b*). Fragment E60B_UCS contains an ultraconserved element that will be described elsewhere and is used here as an anchor point to allow localisation of the conserved/non-conserved E60A element from zebrafish *pax6a* and *pax6b*. A fragment containing the human E60A element was cloned into an hsp68-LacZ reporter construct to generate transient transgenic embryos. The equivalent regions from the zebrafish *pax6a* and *pax6b* loci were cloned into a Tol2 transposon based vector containing a cFos minimal promoter GFP reporter cassette [53] and used to produce transgenic zebrafish.

(B–E) Reporter expression of the human E60A construct was observed in optic cup, telencephalon, and prosomere P3 region of the diencephalon in embryos of E9.5 (B); E10.5 (C and D); and E11.5 (E). Expression (ectopic) in the limbs is ascribed to the genomic integration site of the construct.

(F) A similar GFP reporter expression pattern is seen in transgenic zebrafish embryos with the E60A region from the *pax6a* locus, with GFP expression observed in the optic cup and forebrain.

(G) No reporter expression is seen with the construct containing the E60A region from locus *pax6b*.

(H) Frontal view of a transgenic zebrafish with the *pax6a* E60A element, showing expression in telencephalon and optic cup.

(I) Frontal view of a *pax6b* E60A transgenic fish showing a lack of specific expression. T, telencephalon; OC, optic cup; P3, prosomere P3 of the diencephalon.

doi:10.1371/journal.pgen.0040029.g005

cassette linked to the endogenous promoters in conjunction with a second fragment containing a putative *cis*-element (Figure 6B). Four different combinations of the zebrafish elements were created using the compound enhancer region and the P0 promoter from both *pax6a* and *pax6b* (Figure 6A and 6B). A similar size random sequence DNA was also linked to the zebrafish P0 region as control. YFP expression was assessed by fluorescence and also using RNA *in situ* analysis. The patterns of expression observed for the P/EE(A)-P0(A) (AA) and the P/EE(B)-P0(B) (BB) constructs are illustrated in Figure 6C. At 28 hpf the alkaline phosphatase-labelled YFP *in situ* results reveal expression patterns that recapitulate part of

the endogenous *pax6a* and *pax6b* expression patterns in accordance with their divergent expression in pancreas and MHB. Frequencies of observed expression sites for the four combinations are shown in Figure 6D. All constructs show a high frequency of expression in the telencephalon, including the constructs with a random sequence linked to the P0 promoters, suggesting that telencephalon enhancer activity resides in the P0 promoter regions from both loci. Interestingly, the telencephalic expression frequency appears lower in the presence of P/EE from the *pax6a* locus. In contrast, expression at the MHB and at lower frequency in the pancreas are clearly dependent on the presence of P/EE(B).

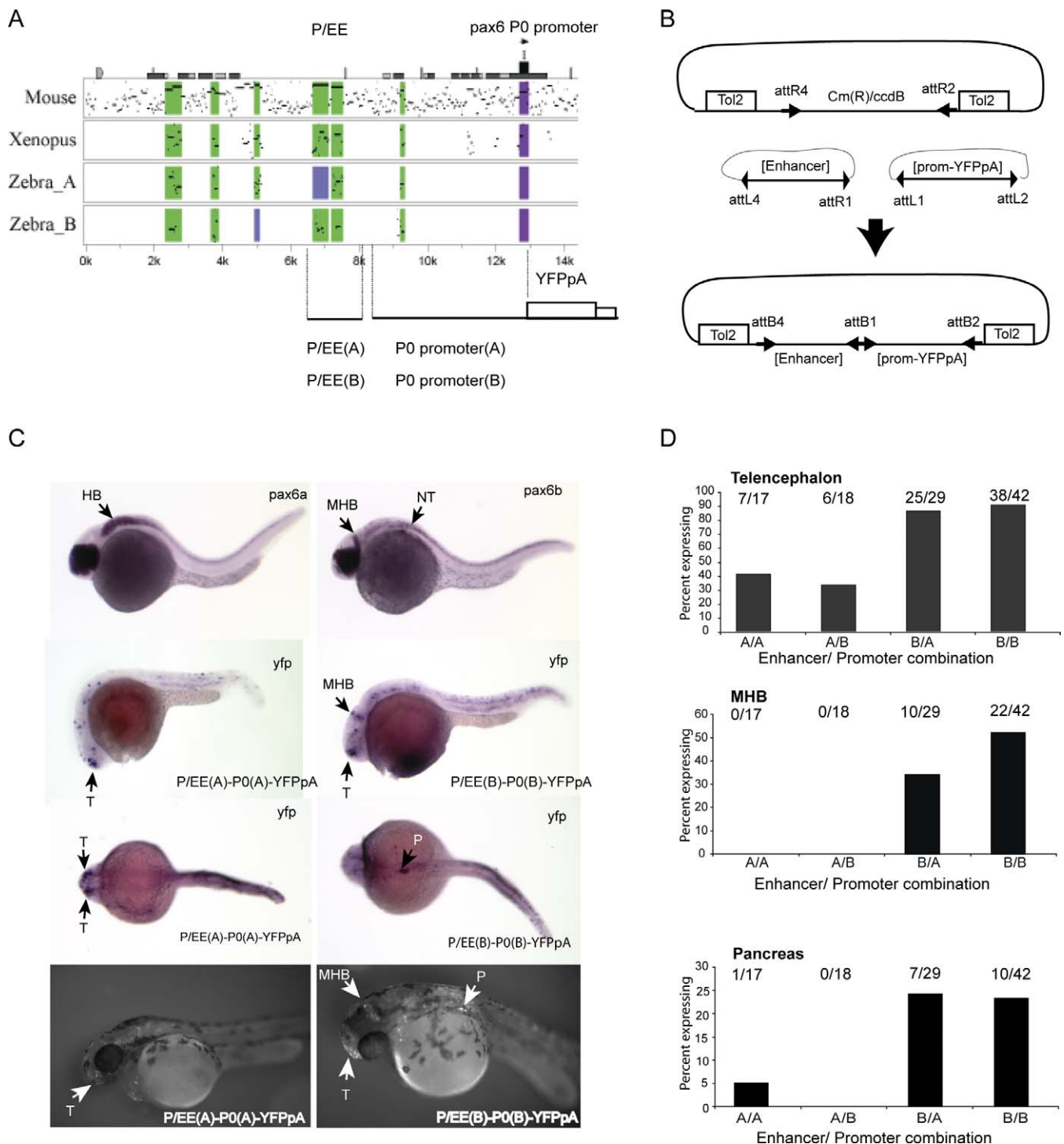


Figure 6. Transgenic Analysis of the Zebrafish P/EE Enhancers from the *pax6a* and *pax6b* Loci in Combination with the *pax6* P0 Promoters (A) PIP plot showing the conservation levels in the *Pax6* upstream regions of mouse, *Xenopus tropicalis*, and the two zebrafish co-orthologues using the human gene as base sequence. The position of the pancreas/ectodermal (P/EE) and the P0 promoter regions used in the construction of the reporter transgenes are shown.

(B) Schematic representation of the Tol2-2way system to create combinatorial constructs for Tol2 mediated zebrafish transgenesis.

(C) Expression analysis of *pax6a* and *pax6b* shown by RNA in situ analysis with alkaline phosphatase staining at 28 hpf (top). YFP expression in P/EE(A)-P0(A)-YFPpA (left) and P/EE(B)-P0(B)-YFPpA (right) transgenic fish at 28 hpf shown, using in situ analysis for YFP mRNA. 56 hpf fluorescence images of YFP expression in transgenic lines for the same two constructs are shown on the bottom row.

(D) Frequency of expression in different sites in transient transgenic zebrafish with the four combinations of P/EE enhancer and P0 promoter from the two *pax6* loci. T, telencephalon; MHB, midbrain hindbrain boundary; HB, hindbrain; P, pancreas; NT, neural tube.

doi:10.1371/journal.pgen.0040029.g006

Constructs with the P/EE(A) element or a random sequence in combination with either promoter show reporter expression only in the telencephalon. Occasionally expression was found in a few isolated cells in the retina with the P/EE containing constructs. More detailed analysis of these expression patterns will be possible in transgene transmitting G1 and subsequent generations, when any transgene mosaicism will have been eliminated.

Concluding Remarks

Gene duplication presents multiple discrete opportunities for differential alteration of gene function in closely related individuals. Divergence at just a few essential loci may lead to reproductive isolation and speciation, as in teleost fish [19–23]. Zebrafish *pax6a* and *pax6b* probably arose through large scale genome duplication, rather than regional duplication and translocation. Their position on different linkage groups, both showing vestiges of conserved synteny with ancestral neighbouring genes, is good evidence for this. Surprisingly, the *pax6* genes, which are strongly dosage sensitive in mammals [30,82–84], have survived duplication, while the ubiquitously expressed flanking genes *ren1* and *elp4*, which do not give rise to a dosage (haploinsufficiency) phenotype in humans [85] or mice [68], have lost their co-orthologues. It is interesting to speculate whether the coordinated loss of tissue-specific regulatory elements in duplicate *pax6* loci is selected for because gene dosage needs to be maintained at the correct level. Further studies in other teleost species may help to resolve this question.

The diverged functions of the *pax6* co-orthologues are unlikely to be due to coding region changes [43], although there are ~20 amino acid differences between them, and also between each zebrafish gene and their single human orthologue. In contrast, the L244P mutation found in the *sri* mutant is deleterious; it shows reduced binding *in vitro* to a canonical paired-type homeodomain target site and its ability to activate both the canonical homeodomain target P3 and the paired domain target CD19 is abolished or severely reduced. In keeping with the expression spectrum of *pax6b*, the mutation gives rise predominantly to a microphthalmic phenotype with variable reduction in lens size and, in extreme cases, disorganisation of the retina (Figure 1A), that may be secondary to extreme lens reduction (cf Ashery-Padan et al, 2000 [80]). The absence of a severe pancreas phenotype is intriguing; our discussions raise various possible explanations, but further analysis is needed for better insight.

Our findings provide strong experimental evidence for the concept of evolutionary divergence by subfunctionalization through differential loss and retention of *cis*-regulatory elements at ancestrally duplicated tissue-specific loci [86]. This has also been termed the “duplication, divergence and complementation” model [87]. For the retention of a full set of functions, the loss of elements needs to proceed in an orderly manner. Interdependent *cis*-elements can only be lost from the same copy of the duplicated gene, otherwise gene activity in certain tissues may be totally lost. This may be the situation where several brain-specific elements, with spatio-temporal overlap in expression pattern, have been lost from *pax6b*. The persistence of *pax6a*-associated regulatory elements, when the exons of the *elp4* gene have disappeared, may be one mechanism for the evolution of gene deserts containing only conserved non-coding sequences [12,88].

Methods

Embryo collection. General maintenance, collection, and staging of zebrafish was carried out according to the Zebrafish Book [89]. The approximate stages are given in hours post fertilization (hpf) at 28 °C and are determined according to morphological criteria [90].

Genetic mapping of the *sri* locus. The zebrafish mutant *sunrise* (allele: *sri*^{tg253a}), produced as part of an ENU mutagenesis screen [27], was provided by the Tübingen zebrafish stockcentre. Homozygous *sri* fish, showing the variable recessive phenotype, were found to be fully viable and fertile. Fish carrying the *sri* mutation on the Tübingen background were crossed to the wildtype strain WIK (provided by Peter Currie). The F1 was raised, and 5 dpf F2 larvae showing the mutant phenotype, and their WT siblings, were collected and fixed in 100% methanol prior to mapping. The *sri* mutation was mapped by linkage analysis with SSLP (simple sequence length polymorphism) markers between embryos showing the homozygous mutant phenotype and their unaffected siblings.

Radiation hybrid mapping. Radiation hybrid (RH) mapping was carried out using the Goodfellow T51 RH panel. To obtain the positions of the tested markers the Children's Hospital Zebrafish Genome Initiative webpage (<http://zfrhmaps.tch.harvard.edu/ZonRHmapper/instantMapping.htm>) was used. Intronic primers flanking exons 1 and 6 of *pax6b* were used to map this strong candidate gene onto the T51 RH map that was also used as a reference for the markers used in the genetic mapping. The map position of *pax6b* was closest to the marker *fc33g08*, which is now known to be part of the *pax6b* coding region in NM_131641 at <http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&val=18859210>.

Mutation analysis using RT-PCR and sequencing. *pax6b* cDNA was derived by RT-PCR from *sri* and WT embryos using Superscript II (Gibco). Sequencing revealed four changes identified in the *sri* coding sequence compared to the published *pax6b* cDNA [43] (Genbank NM_131641). Three of the four nucleotide changes were synonymous (429 A>C, 989 G>A, 1304 A>C), and also found to be present in the wildtype sequence of the Tü strain, in which the *sri* mutation was generated [15], as well as in the public genomic sequence in BAC DKEY-157G7. The fourth, a T to C transition at position 991 in the cDNA is present only in *sri* DNA and gives rise to a leucine to proline missense mutation in the homeodomain. The 991 T>C mutation creates a *DdeI* digestion site that was used for further genotyping (Figure S2).

Functional studies of the *sri* mutant protein. Luciferase assays were carried out as follows. Mammalian expression vectors, pCMV-ZfPax6.1 (*pax6a*), pCMV-ZfPax6.2 (*pax6b*) and pCMV-ZfPax6.2-*sri* (*pax6b-sri*), respectively containing the two wild type or L244P mutant full-length cDNA clone, were used in co-transfections into HeLa cells with the luciferase reporter plasmids 3XP3-Luc and 3XCD19-Luc, containing respectively three P3 homeodomain binding sites [45] or three CD19 paired domain binding sites [46] linked 5' to the basal promoter-luciferase cassette. HeLa cells were cultured in DMEM medium with 10% calf serum (Gibco, Milan-Italy) and transfected using calcium phosphate co-precipitation [91]. Cells were plated at 6×10^5 cells/100 mm culture dish 20h prior to transfection. The plasmids used were: pCMV-ZfPax6b and pCMV-ZfPax6b:*sri* 1 µg; 3XP3-Luc: 9 µg; RSV-CAT: 2 µg. RSV-CAT contains the RSV promoter 5' to the choramphenicol-acetyl-transferase (CAT) gene, used to normalize transfection efficiency. Cells were harvested 42–44 h after transfection, and extracts prepared by freeze-thawing. CAT activity was measured by ELISA (Amersham). Luciferase activity by chemiluminescence. The data presented are mean values plus/minus standard deviations of a representative experiment performed in triplicate (three independent transfections).

Electrophoretic mobility shift assays (EMSA) were conducted as follows. cDNA sequences encoding either wild-type ZF *pax6b* or mutant L244P homeodomain (from aminoacid 228 to amino acid 291 of the full-length protein) were cloned in the pT7.7 bacterial expression vector, checked by nucleotide sequencing, expressed in BL21 cells, and partially purified. ³²P end-labeled double-stranded oligodeoxynucleotides P2 and P3 containing high-affinity HD-binding sequences

P2: 5'-TCGAGGGCATCAGGATGCTAATTGATTAGCATCC-GATCGG -3';

P3: 5'-TCGAGGGCATCAGGATGCTAATTGGATTAGCATCC-GATCGG -3'

were used in gel-retardation assays [47,92]. HD-containing protein at a concentration of 50, 100 and 300 ng/sample was incubated for 30 minutes at room temperature with probe DNA (5 mM in a buffer containing 20 mM Tris-HCl pH 7.6, 75 mM KCl, 0.25 mg/ml bovine serum albumin (BSA), 5mM dithiothreitol (DTT), 12.5 µg/ml calf thymus DNA, 10% glycerol). Protein-bound and free oligos were

separated on native 7.5 % polyacrylamide gel, run in $0.5 \times$ TBE for 1.5 hours at 4 °C. Gels were fixed and autoradiographed. Electrophoretic mobility shift assays were carried out for wild type and mutant homeodomain protein fragments on the P2 and P3 homeodomain binding sites.

For histological analyses, *sri* mutant and age-matched sibling larvae were fixed in 4 % paraformaldehyde in PBS, dehydrated in a standard ethanol series, embedded in Technovit (Heraeus Kulzer, Germany), sectioned and stained with 0.5% toluidine blue in 1% borax buffer.

Immunohistochemistry was performed using standard methods and suitably cross-reacting antibodies. Mouse monoclonal anti-porcine glucagon antibody (Sigma G2654) and guinea-pig polyclonal anti-porcine insulin antibody (DAKO A0564) were used with fluorescently labelled second antibodies. Individuals studied were genotyped, as described.

Comparative sequence analysis was carried by through PIP plots, which were made using the programme PIPmaker (<http://bio.cse.psu.edu/pipmaker/>) [75]. Evolutionary sequence comparison of element E60+ was prepared using the VISTA program [78] with a window size of 50 bp and a minimal sequence identity of 70% using sequences derived from Z83001 (human), AL512589 (mouse), AL021531 (pufferfish), and AL929172, BX004784 (zebrafish *pax6a*), AC127461, BX000453 (zebrafish *pax6b*). *Xenopus* sequence was obtained from *Xenopus tropicalis* genome assembly 3.0 from the JGI.

Expression studies of *pax6a*, *pax6b*, and YFP by RNA in situ analysis in zebrafish. Preliminary information is in the public domain [49] and can be accessed directly at the ZFIN database (<http://zfin.org/cgi-bin/webdriver?Mval=aa-xpatselct.apg>) (Accessed: 16 January 2008).

Whole-mount *in situ* hybridization reactions were carried out according to published protocols [93]. The cDNA constructs were prepared through cloning of sequence verified RT-PCR products using RNA prepared from wild type and *sunrise* mutant zebrafish into the eukaryotic expression vector pJG1 (gift from Jacqueline Guy). Riboprobes for *pax6a* and *pax6b* were transcribed from PCR templates amplified from plasmid DNA using the following oligonucleotides:

*pax6a*fw:5'-CCCAATACTGGCCAGACTA-3'
*pax6a*Trv:5'-TAAGCTTTAATACGACTCACTATAGGGAGA-GAAGTGGCACTATCCCCGTA-3'
*pax6b*fw:5'-GAGCAAGATTCTGGGGAGGT-3'
*pax6b*Trv:5'-TAAGCTTTAATACGACTCACTATAGGGA-GAGCTCGGTATGTTATCGTTGG-3'
*eYFP*fw: 5'-CACATGAAGCAGCAGCACTT-3'
*eYFP*Trv: 5'-TAAGCTTTAATACGACTCACTATAGGGA-GAAGTTCACCTTGATGCGCTTC-3'

Sections of whole-mount *in situ* stained embryos were obtained by wax embedding and sectioning at 7 μ m.

Expression analysis of adult zebrafish tissues by RT-PCR. RNA was made from carefully dissected adult (age >6 months) pancreas from both wildtype and homozygous mutant *sunrise* fish with Trizol (Sigma), and reverse transcribed with AMV reverse transcriptase (Roche) and random hexamer primers. Wild type eye tissue was processed alongside as control tissue expressing both *pax6a* and *pax6b*. The resulting cDNA was PCR amplified with primer pairs specific for *pax6a*, *pax6b* or *rcn1*, a ubiquitously expressed calcium binding protein. Primers used: *pax6a* (5'-cagtacaagaggagtgctcc-3' and 5'-ctcaccgctcctgctgactgt-3'),

pax6b (5'-CAGTACAAGAGGGAGTGTCC-3' and 5'-GTTTTCAC-CACCGTTGTCTCTGT-3') and *rcn1* (5'-TTCCACGAGATGAACGCAGGT-3' and 5'-TTCATCTTCATCGACATGACCATC-3').

Antisense morpholino oligonucleotide injections. The *pax6b* antisense morpholino oligonucleotides (MO) (Gene Tools, LLC) were directed against the 5' sequence near the start of translation [94]. The MO used were:

pax6a v1-MO:5-CGTTTACAATAACGAGAAGACTCTG-3 *pax6a* v2-MO:5-AGTTCCAACAGCCTTTGTATCCTCG-3 *pax6b* v1-MO:5-GCCTGAGCCCTTCCGAGCAAATCAG-3 *pax6b* v2-MO: 5-TGTTATTCTTTTGGGCACTTCTGC-3

They were injected in a volume of 1.4 nl through the chorion of 1 to 2-cell stage embryos to deliver a mass of 3.5 ng. For each MO, at least 100 embryos were injected and the phenotype was of consistent severity and fully penetrant. Co-injection of *pax6a* and *pax6b* morpholinos was also tried.

RNA injections for phenotype rescue. The wild-type and *sri* mutant *pax6b* injection construct was cloned into pJG1 from fish-derived cDNAs as described above. The capped mRNA was synthesized using the mMessage mMachine kit T7 (Ambion) and purified using the RNeasy Mini kit (Qiagen). The RNAs were injected into 1 to 2-cell stage embryos as described previously.

Reporter transgenesis in zebrafish using the Tol2 system. The E60A_cfosEGFP and E60B_cfosEGFP (Figures 6 and S7C) were generated according to published protocols [53], using the following oligonucleotides:

pax6a E60afw: 5'-GGGGACAAGTTTGTACAAAAAAGCAGGC-TAACAAATGTCACAGACTGCAAGA-3'
pax6a E60arv: 5'-GGGGACCACTTTGTACAAAGAAAGCTGGGTG-CGCGCCTACAACCACGTTTCTACCAGC-3'

pax6b E60afw: 5'-GGGGACAAGTTTGTACAAAAAAGCAGGC-TACTGGAAAATATACAGAGGGC-3'

pax6b E60arv:5'-GGGGACCACTTTGTACAAAGAAAGCTGGGTG-CGCGCCATTCATCCGTCTGTCCATCCA-3' Fish were injected as described for RNA injections.

Tol2-2way system to create combinatorial constructs for Tol2 mediated zebrafish transgenesis. A vector was created in which an AttR4-R2 Gateway 2-way recombination cassette (Invitrogen) was flanked by *cis*-sequences required for Tol2 mediated transposition based on the medaka Tol2 transposon [53] generating the pTol2-2way destination vector. Putative enhancer regions are cloned into an attP4-P1r plasmid and combined with a promoter-reporter cassette containing attL1-L2 plasmid and the pTol2-2way vector to generate the microinjection construct. Various combinations of those two constructs can be recombined into the pTol2-2way destination vector. The entry clones for the P/EE elements or the random sequence (RS; a non-conserved region in a gene desert on human chromosome 17) were made using the following primers:

attB4-P/EE(a) primer: 5'-AGGGGACAACCTTTGTATAGAAAAG-TTGGCGCGCCTCTCCGCCAATGAATCTGCA-3'
 attB1r-P/EE(a) primer: 5'-AGGGGACTGCTTTTTTTGTACAAACT-TGTAGTTGGATATTACAGGCAGT-3'
 attB4-P/EE(b) primer: 5'-AGGGGACAACCTTTGTATAGAAAAG-TTGGCGCGCCATGAACAGACAGATATGGCA-3'
 attB1r-P/EE(b) primer: 5'-AGGGGACTGCTTTTTTTGTACAAACT-TGAACGTCTCGACCATGCAGGCT-3'
 attB4-RS primer: 5'-AGGGGACAACCTTTGTATAGAAAAGTTGG-CGCGCCATTCATTAGATCAATCCCT-3'
 attB1r-RS primer: 5'-AGGGGACTGCTTTTTTTGTACAAACTTGA-TTCTTGCTGGGTAGGGTACA-3'

The promoter-reporter constructs were made by cloning of the P0 promoter regions from the zebrafish *pax6a* (2 kb) and *pax6b* (2.5 kb) loci in front of a promoterless YFPpolyA cassette in an attL1-L2 containing pDONR vector, using an NcoI restriction site and the following primer sets:

pax6a P0 promoter: 5'-AACCATGGACTGCCTGTAATATCCAAC-TAC-3' and 5'-ACCATGGTGGCGGCAGTCCAACAAGGGAACCTC-GA-3'.

pax6b P0 promoter: 5'-ACCATGGATCTGCGACATGCTCATGT-GA-3' and 5'-ACCATGGCGCGCCGCTGCTGTTCAAGCCCTT-3'.

Analysis of E60 reporter constructs in transgenic mice. The E60A-Z reporter constructs are based on a modified p610 vector containing a hsp68-*LacZ* cassette [6]. A 2.3 kb fragment containing the E60+ conserved region was PCR amplified from human cosmid G0453 (EMBL Z83001) using high fidelity polymerase (Bio-X-Act) with primer pair (5'-TGTCGACCAATGCAGCACCAAGTGTATGC-3') (5'-AGTCGACAAAAGATAAGTAAGCTCAGATGT-3'), and cloned into the SalI site of the p610+ vector to generate E60+Z. The E60A-Z microinjection fragment was generated from this construct using the Asp718I restriction site located between the two subregions of the full-length E60+ fragment. All the expressions patterns were photographed on a Leica MZ FLIII Microscope fitted with a Hamamatsu Orca-ER digital camera.

Supporting Information

Figure S1. Identification and Validation of the *sri* Mutation

(A) Mapping of the *sri* mutation to the *pax6b* region on LG7. Markers used in the region are shown for a Radiation Hybrid (RH) map and by linkage.

(B) Alignment of the amino acid sequences for the PAX6 homeodomain with Pax6 proteins from multiple other species and with other human homeodomains from multiple genes with paired-type homeoboxes. The alignment shows (arrowed) that the L244P change has arisen in a totally conserved leucine residue.

(C) Documentation of the nucleotide differences (red boxes) between deposited genomic and cDNA sequences for *pax6b* and showing the position of the proposed ENU-induced CTT to CCT nucleotide change leading to the missense mutation. All the other nucleotide variants (strain differences) do not lead to protein residue changes.

Found at doi:10.1371/journal.pgen.0040029.sg001 (323 KB PPT).

Figure S2. Genotyping the Offspring of Two Heterozygous Parents by Restriction Analysis Reveals the Expected Numbers of Each Genotype
Found at doi:10.1371/journal.pgen.0040029.sg002 (197 KB PPT).

Figure S3. RT-PCR Analysis Shows That Only *pax6b* Is Expressed in the Pancreas of Both Wild Type and *sri/sri* Homozygous Adult Fish. In Contrast, Both *pax6a* and *pax6b* Are Expressed in Adult Eye
Found at doi:10.1371/journal.pgen.0040029.sg003 (199 KB PPT).

Figure S4. BAC Assembly of the LG25 *pax6a* Region, Showing the Syntenic Conservation with *wt1a*

The other two immediate *pax6* flanking genes normally seen in vertebrates, *elpl4* and *rcn1*, are however absent from this contig.
Found at doi:10.1371/journal.pgen.0040029.sg004 (496 KB PPT).

Figure S5. BAC Assembly of the LG7 *pax6b* Region, Showing the Presence of *elpl4* and *rcn1*
Found at doi:10.1371/journal.pgen.0040029.sg005 (56 KB PPT).

Figure S6. Sequence Comparison around *pax6* by PIP Plot, with *Fugu* as the Base Sequence, Showing That *elpl4* Exons Are Present Downstream of Zebrafish *pax6b*, but Absent from *pax6a*
Found at doi:10.1371/journal.pgen.0040029.sg006 (199 KB PPT).

Figure S7. Sequence Comparison around *pax6* by PIP Plot, with Chick as the Base Sequence Comparing to Human, Mouse, *Fugu* and *pax6a* and *pax6b* from Zebrafish

The enhancer elements discussed in this paper are highlighted with red boxes (P/EE, E60, and CE2). A blue box draws attention to the greater conservation of regulatory elements downstream of *pax6a* than *pax6b*, which is noteworthy since *pax6a* has lost the *elpl4* exons, while *pax6b* has retained them.

Found at doi:10.1371/journal.pgen.0040029.sg007 (315 KB PPT).

Figure S8. Details of the E60 Elements in Mouse and Zebrafish Transgenic Studies

(A) Mouse human and zebrafish comparisons for the E60A region.
(B) E60A sequence from the *pax6b* region, with no observable conserved sequence to *pax6a* E60A or to other vertebrate homologues, was tested in the zebrafish transgenics.
(C) The making of Tol2 constructs for efficient zebrafish transgenesis.
Found at doi:10.1371/journal.pgen.0040029.sg008 (163 KB PPT).

Figure S9. Summary of Conserved Elements within and Flanking the

PAX6 Gene, Most with Proven Tissues-Specific, *Pax6* Compatible Enhancer Function in the Mouse.

Found at doi:10.1371/journal.pgen.0040029.sg009 (63 KB PPT).

Accession Numbers

The RefSeq (National Center for Biotechnology Information [NCBI]) (<http://www.ncbi.nlm.nih.gov/RefSeq>) accession numbers for the genes discussed in this paper are for human: *PAX6* (NM_000280), *NP_001595*; *ELP4* (NM_019040); *RCN1* (NM_002901); *WT1* (NM_024425); mouse: *Pax6* (NM_013627); *Elpl4* (NM_023876); *Wt1* (NM_144783); Zebrafish: *pax6a* (NM_131304); *pax6b* (NM_131641); *elpl4* (NM_001017638); *rcn1* (predicted) (XM_686046); *wt1a* (NM_131046); *wt1b* (NM_001039634).

Diseases, etc.: human aniridia (MIM 106210); mouse small eye (MGI:1856155); zebrafish sunrise (ZDB-GENE-070117-2413).

Acknowledgments

We thank Robert Geisler for the use of his mapping set-up; the Thisse lab for the gift of the *pax6a* and *pax6b* cDNA constructs; Koichi Kawakami for the pCS-Tp plasmid for Tol2 transgenesis; and A. McCallion and S. Fisher for the pGW_cfosEGFP reporter vector. John Maule and Peter Currie and colleagues provided invaluable help and advice for the zebrafish work; Paul Perry helped with imaging; Brendan Doe and colleagues gave assistance with mouse transgenic studies. We appreciated the opportunity for stimulating discussions with Nick Hastie and David FitzPatrick.

Author contributions. DAK, RMB, VvH, and PC conceived and designed the experiments and analyzed the data with significant input by PLY. RMB carried out the mating and phenotype selection and confirmations for the mapping. DAK and PC developed the transgenic strategies and constructs and made and analyzed mouse and fish transgenics with help from AS. PC carried out mRNA rescue and morpholino experiments. PG and PC provided bioinformatics contributions. RD supplied the sunrise mutant and worked with HBS on the fish mapping. RD and PLY analyzed the fish eye phenotype. AMH analyzed the pancreas. GD conceived and carried out the mutant protein functional analyses. VvH wrote the paper.

Funding. The following support is gratefully acknowledged: VvH lab is funded by the Medical Research Council, UK; AMH is funded by Fight for Sight; PC is funded by Marie Curie International Fellowship FP6-MEIF-CT-2005-025389; and RD is funded by EMBO Long-term Fellowship.

Competing interests. The authors have declared that no competing interests exist.

References

- Levine M, Tjian R (2003) Transcription regulation and animal diversity. *Nature* 424: 147–151.
- Kleinjan DA, van Heyningen V (2005) Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet* 76: 8–32.
- Ahituv N, Rubin EM, Nobrega MA (2004) Exploiting human–fish genome comparisons for deciphering gene regulation. *Hum Mol Genet* 13: R261–R266.
- Plessy C, Dickmeis T, Chalmel F, Strahle U (2005) Enhancer sequence conservation between vertebrates is favoured in developmental regulator genes. *Trends Genet* 21: 207–210.
- Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, et al. (2006) In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444: 499–502.
- Kleinjan DA, Seawright A, Schedl A, Quinlan RA, Danes S, et al. (2001) Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. *Hum Mol Genet* 10: 2049–2059.
- Griffin C, Kleinjan DA, Doe B, van Heyningen V (2002) New 3' elements control Pax6 expression in the developing pretectum, neural retina and olfactory region. *Mech Dev* 112: 89–100.
- Lettice LA, Heaney SJ, Purdie LA, Li L, de Beer P, et al. (2003) A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet* 12: 1725–1735.
- Kleinjan DA, Seawright A, Childs AJ, van Heyningen V (2004) Conserved elements in Pax6 intron 7 involved in (auto)regulation and alternative transcription. *Dev Biol* 265: 462–477.
- Woolfe A, Goodson M, Goode DK, Snell P, McEwen GK, et al. (2005) Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol* 3: e7. doi:10.1371/journal.pbio.0030007
- Jeong Y, El-Jaick K, Roessler E, Muenke M, Epstein DJ (2006) A functional screen for sonic hedgehog regulatory elements across a 1 Mb interval identifies long-range ventral forebrain enhancers. *Development* 133: 761–772.
- Kikuta H, Laplante M, Navratilova P, Komisarczuk AZ, Engstrom PG, et al. (2007) Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. *Genome Res* 17: 545–555.
- Lupski JR, Stankiewicz P (2005) Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet* 1: e49. doi:10.1371/journal.pgen.0010049
- Loots GG, Kneissel M, Keller H, Baptist M, Chang J, et al. (2005) Genomic deletion of a long-range bone enhancer misregulates sclerostin in Van Buchem disease. *Genome Res* 15: 928–935.
- Mohrs M, Blankespoor CM, Wang ZE, Loots GG, Afzal V, et al. (2001) Deletion of a coordinate regulator of type 2 cytokine expression in mice. *Nat Immunol* 2: 842–847.
- Zuniga A, Michos O, Spitz F, Haramis AP, Panman L, et al. (2004) Mouse limb deformity mutations disrupt a global control region within the large regulatory landscape required for Gremlin expression. *Genes Dev* 18: 1553–1564.
- Tarchini B, Huynh TH, Cox GA, Duboule D (2005) HoxD cluster scanning deletions identify multiple defects leading to paralysis in the mouse mutant Ironside. *Genes Dev* 19: 2862–2876.
- Wittkopp PJ, Haerum BK, Clark AG (2004) Evolutionary changes in cis and trans gene regulation. *Nature* 430: 85–88.
- Force A, Lynch M, Pickett FB, Amores A, Yan YL, et al. (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151: 1531–1545.
- Postlethwait J, Amores A, Cresko W, Singer A, Yan YL (2004) Subfunction partitioning, the teleost radiation and the annotation of the human genome. *Trends Genet* 20: 481–490.

21. Volff JN (2005) Genome evolution and biodiversity in teleost fish. *Heredity* 94: 280–294.
22. Venkatesh B, Kirkness EF, Loh YH, Halpern AL, Lee AP, et al. (2006) Ancient noncoding elements conserved in the human genome. *Science* 314: 1892.
23. Semon M, Wolfe KH (2007) Reciprocal gene loss between Tetraodon and zebrafish after whole genome duplication in their ancestor. *Trends Genet* 23: 108–112.
24. Woolfe A, Elgar G (2007) Comparative genomics using Fugu reveals insights into regulatory subfunctionalization. *Genome Biol* 8: R53.
25. Loosli F, Staub W, Finger-Baier KC, Ober EA, Verkade H, et al. (2003) Loss of eyes in zebrafish caused by mutation of *chokh/rx3*. *EMBO Rep* 4: 894–899.
26. Yeyati PL, Bancewicz RM, Maule J, van Heyningen V (2007) Hsp90 selectively modulates phenotype in vertebrate development. *PLoS Genet* 3: e43. doi:10.1371/journal.pgen.0030043
27. Heisenberg CP, Brand M, Jiang YJ, Warga RM, Beuchle D, et al. (1996) Genes involved in forebrain development in the zebrafish, *Danio rerio*. *Development* 123: 191–203.
28. Callaerts P, Munoz-Marmol AM, Glardon S, Castillo E, Sun H, et al. (1999) Isolation and expression of a pax-6 gene in the regenerating and intact planarian *Dugesia (G.) tigrina*. *Proc Natl Acad Sci U S A* 96: 558–563.
29. van Heyningen V, Williamson KA (2002) PAX6 in sensory development. *Hum Mol Genet* 11: 1161–1167.
30. Hill RE, Favor J, Hogan BM, Ton CCT, Saunders GF, et al. (1991) Mouse small eye results from mutations in a paired-like homeobox- containing gene. *Nature* 354: 522–525.
31. Glaser T, Jepeal L, Edwards JG, Young SR, Favor J, et al. (1994) PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. *Nat Genet* 7: 463–471.
32. Quiring R, Walldorf U, Kloter U, Gehring WJ (1994) Homology of the eyeless gene of *Drosophila* to the Small eye gene in mice and Aniridia in humans. *Science* 265: 785–789.
33. Callaerts P, Leng S, Clements J, Benassayag C, Cribbs D, et al. (2001) *Drosophila* Pax-6/eyeless is essential for normal adult brain structure and function. *J Neurobiol* 46: 73–88.
34. Sisodiya SM, Free SL, Williamson KA, Mitchell TN, Willis C, et al. (2001) PAX6 haploinsufficiency causes cerebral malformation and olfactory dysfunction in humans. *Nat Genet* 28: 214–216.
35. Estivill-Torres G, Vitalis T, Fernandez-Llebreg P, Price DJ (2001) The transcription factor Pax6 is required for development of the diencephalic dorsal midline secretory radial glia that form the subcommissural organ. *Mech Dev* 109: 215–224.
36. Ashery-Padan R, Zhou X, Marquardt T, Herrera P, Toubé L, et al. (2004) Conditional inactivation of Pax6 in the pancreas causes early onset of diabetes. *Dev Biol* 269: 479–488.
37. Fantes J, Redeker B, Breen M, Boyle S, Brown J, et al. (1995) Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Hum Mol Genet* 4: 415–422.
38. Kleinjan DA, Seawright A, Mella S, Carr CB, Tyas DA, et al. (2006) Long-range downstream enhancers are essential for Pax6 expression. *Dev Biol* 299: 563–581.
39. Plaza S, Saule S, Dozier C (1999) High conservation of cis-regulatory elements between quail and human for the Pax-6 gene. *Dev Genes Evol* 209: 165–173.
40. Visel A, Bristow J, Pennacchio LA (2007) Enhancer identification through comparative genomics. *Semin Cell Dev Biol* 18: 140–152.
41. Geisler R, Rauch GJ, Baier H, van Bebber F, Brobeta L, et al. (1999) A radiation hybrid map of the zebrafish genome. *Nat Genet* 23: 86–89.
42. Geisler R (2002) Mapping and cloning. Nusslein-Volhard C, Dahm R, editors. *Zebrafish: a practical approach*. Oxford: Oxford University Press. pp. 175–212.
43. Nornes S, Clarkson M, Mikkola I, Pedersen M, Bardsley A, et al. (1998) Zebrafish contains two pax6 genes involved in eye development. *Mech Dev* 77: 185–196.
44. Banerjee-Basu S, Landsman D, Baxevanis AD (1999) Threading analysis of prospero-type homeodomains. *In Silico Biol* 1: 163–173.
45. Wilson D, Sheng G, Lecuit T, Dostatni N, Desplan C (1993) Cooperative dimerization of paired class homeo domains on DNA. *Genes Dev* 7: 2120–2134.
46. Epstein JA, Glaser T, Cai J, Jepeal L, Walton DS, et al. (1994) Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. *Genes Dev* 8: 2022–2034.
47. Singh S, Stellrecht CM, Tang HK, Saunders GF (2000) Modulation of PAX6 homeodomain function by the paired domain. *J Biol Chem* 275: 17306–17313.
48. Mishra R, Gorlov IP, Chao LY, Singh S, Saunders GF (2002) PAX6, paired domain influences sequence recognition by the homeodomain. *J Biol Chem* 277: 49488–49494.
49. Thisse B, Thisse C (2004) Fast release clones: a high throughput expression analysis. Available at: <http://zfinfo.org/cgi-bin/webdriver?MVal=aa-pubview2.apg&OID=ZDB-PUB-040907-1>. Accessed: 16 January 2008.
50. Hanson I, Churchill A, Love J, Axton R, Moore T, et al. (1999) Missense mutations in the most ancient residues of the PAX6 paired domain underlie a spectrum of human congenital eye malformations. *Hum Mol Genet* 8: 165–172.
51. Tzoulaki I, White IM, Hanson IM (2005) PAX6 mutations: genotype-phenotype correlations. *BMC Genet* 6: 27.
52. Nakano Y, Kim HR, Kawakami A, Roy S, Schier AF, et al. (2004) Inactivation of dispatched 1 by the chameleon mutation disrupts Hedgehog signalling in the zebrafish embryo. *Dev Biol* 269: 381–392.
53. Fisher S, Grice EA, Vinton RM, Bessling SL, Urasaki A, et al. (2006) Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish. *Nat Protoc* 1: 1297–1305.
54. Kim J, Lauderdale JD (2006) Analysis of Pax6 expression using a BAC transgene reveals the presence of a paired-less isoform of Pax6 in the eye and olfactory bulb. *Dev Biol* 292: 486–505.
55. Manuel M, Georgala PA, Carr CB, Chanas S, Kleinjan DA, et al. (2007) Controlled overexpression of Pax6 in vivo negatively autoregulates the Pax6 locus, causing cell-autonomous defects of late cortical progenitor proliferation with little effect on cortical arealization. *Development* 134: 545–555.
56. Favor J, Peters H, Hermann T, Schmahl W, Chatterjee B, et al. (2001) Molecular characterization of Pax6(2Neu) through Pax6(10Neu): an extension of the Pax6 allelic series and the identification of two possible hypomorph alleles in the mouse *Mus musculus*. *Genetics* 159: 1689–1700.
57. Thauung C, West K, Clark BJ, McKie L, Morgan JE, et al. (2002) Novel ENU-induced eye mutations in the mouse: models for human eye disease. *Hum Mol Genet* 11: 755–767.
58. Graw J, Loster J, Puk O, Munster D, Haubst N, et al. (2005) Three novel Pax6 alleles in the mouse leading to the same small-eye phenotype caused by different consequences at target promoters. *Invest Ophthalmol Vis Sci* 46: 4671–4683.
59. Morrison D, FitzPatrick D, Hanson I, Williamson K, van Heyningen V, et al. (2002) National study of microphthalmia, anophthalmia, and coloboma (MAC) in Scotland: investigation of genetic aetiology. *J Med Genet* 39: 16–22.
60. Haubst N, Berger J, Radjendirane V, Graw J, Favor J, et al. (2004) Molecular dissection of Pax6 function: the specific roles of the paired domain and homeodomain in brain development. *Development* 131: 6131–6140.
61. Ritz-Laser B, Estreicher A, Klages N, Saule S, Philippe J (1999) Pax-6 and Cdx-2/3 interact to activate glucagon gene expression on the G1 control element. *J Biol Chem* 274: 4124–4132.
62. Beimesche S, Neubauer A, Herzig S, Grzeskowiak R, Diedrich T, et al. (1999) Tissue-specific transcriptional activity of a pancreatic islet cell-specific enhancer sequence/Pax6-binding site determined in normal adult tissues in vivo using transgenic mice. *Mol Endocrinol* 13: 718–728.
63. Chauhan BK, Yang Y, Cveklova K, Cvek A (2004) Functional properties of natural human PAX6 and PAX6(5a) mutants. *Invest Ophthalmol Vis Sci* 45: 385–392.
64. Gronskov K, Rosenberg T, Sand A, Brondum-Nielsen K (1999) Mutational analysis of PAX6: 16 novel mutations including 5 missense mutations with a mild aniridia phenotype. *Eur J Hum Genet* 7: 274–286.
65. Kammandel B, Chowdhury K, Stoykova A, Aparicio S, Brenner S, et al. (1999) Distinct cis-essential modules direct the time-space pattern of the Pax6 gene activity. *Dev Biol* 205: 79–97.
66. Kent J, Lee M, Schedl A, Boyle S, Fantes J, et al. (1997) The reticulocalbin gene maps to the WAGR region in human and to the Small eye Harwell deletion in mouse. *Genomics* 42: 260–267.
67. Miles C, Elgar G, Coles E, Kleinjan DJ, van Heyningen V, et al. (1998) Complete sequencing of the fugu WAGR region from WT1 to PAX6: dramatic compaction and conservation of synteny with human chromosome 11p13. *Proc Natl Acad Sci U S A* 95: 13068–13072.
68. Kleinjan DA, Seawright A, Elgar G, van Heyningen V (2002) Characterization of a novel gene adjacent to PAX6, revealing synteny conservation with functional significance. *Mamm Genome* 13: 102–107.
69. Williams SC, Altmann CR, Chow RL, Hemmati-Brivanlou A, Lang RA (1998) A highly conserved lens transcriptional control element from the Pax-6 gene. *Mech Dev* 73: 225–229.
70. Plaza S, Dozier C, Langlois MC, Saule S (1995) Identification and characterization of a neuroretina-specific enhancer element in the quail Pax-6 (Pax-QNR) gene. *Mol Cell Biol* 15: 892–903.
71. Xu PX, Zhang X, Heaney S, Yoon A, Michelson AM, et al. (1999) Regulation of Pax6 expression is conserved between mice and flies. *Development* 126: 383–395.
72. Taylor JS, Braasch I, Frickey T, Meyer A, Van de PY (2003) Genome duplication, a trait shared by 22000 species of ray-finned fish. *Genome Res* 13: 382–390.
73. Woods IG, Wilson C, Friedlander B, Chang P, Reyes DK, et al. (2005) The zebrafish gene map defines ancestral vertebrate chromosomes. *Genome Res* 15: 1307–1314.
74. Bollig F, Mehringer R, Perner B, Hartung C, Schafer M, et al. (2006) Identification and comparative expression analysis of a second wt1 gene in zebrafish. *Dev Dyn* 235: 554–561.
75. Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, et al. (2000) PipMaker—a web server for aligning two genomic DNA sequences. *Genome Res* 10: 577–586.
76. Hughes JR, Cheng JF, Ventress N, Prabhakar S, Clark K, et al. (2005) Annotation of cis-regulatory elements by identification, subclassification,

- and functional assessment of multispecies conserved sequences. *Proc Natl Acad Sci U S A* 102: 9830–9835.
77. Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, et al. (2004) Ultraconserved elements in the human genome. *Science* 304: 1321–1325.
 78. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I (2004) VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 32: W273–W279.
 79. Fisher S, Grice EA, Vinton RM, Bessling SL, McCallion AS (2006) Conservation of RET regulatory function from human to zebrafish without sequence similarity. *Science* 312: 276–279.
 80. Ashery-Padan R, Marquardt T, Zhou X, Gruss P (2000) Pax6 activity in the lens primordium is required for lens formation and for correct placement of a single retina in the eye. *Genes Dev* 14: 2701–2711.
 81. Dimanlig PV, Faber SC, Auerbach W, Makarenkova HP, Lang RA (2001) The upstream ectoderm enhancer in Pax6 has an important role in lens induction. *Development* 128: 4415–4424.
 82. Glaser T, Walton DS, Maas RL (1992) Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene. *Nat Genet* 2: 232–239.
 83. Schedl A, Ross A, Lee M, Engelkamp D, Rashbass P, et al. (1996) Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. *Cell* 86: 71–82.
 84. Davis-Silberman N, Kalich T, Oron-Karni V, Marquardt T, Kroeber M, et al. (2005) Genetic dissection of Pax6 dosage requirements in the developing mouse eye. *Hum Mol Genet* 14: 2265–2276.
 85. Crolla JA, van Heyningen V (2002) Frequent chromosome aberrations revealed by molecular cytogenetic studies in patients with aniridia. *Am J Hum Genet* 71: 1138–1149.
 86. McEwen GK, Woolfe A, Goode D, Vavouri T, Callaway H, et al. (2006) Ancient duplicated conserved noncoding elements in vertebrates: a genomic and functional analysis. *Genome Res* 16: 451–465.
 87. Hurley I, Hale ME, Prince VE (2005) Duplication events and the evolution of segmental identity. *Evol Dev* 7: 556–567.
 88. Taylor J (2005) Clues to function in gene deserts. *Trends Biotechnol* 23: 269–271.
 89. Westerfield M (1995) The zebrafish book. A guide for the laboratory use of zebrafish (*Danio rerio*). Eugene (Oregon): University of Oregon Press.
 90. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF (1995) Stages of embryonic development of the zebrafish. *Dev Dyn* 203: 253–310.
 91. Francis-Lang H, Zannini M, De FM, Berlingieri MT, Fusco A, et al. (1992) Multiple mechanisms of interference between transformation and differentiation in thyroid cells. *Mol Cell Biol* 12: 5793–5800.
 92. Damante G, Pellizzari L, Esposito G, Fogolari F, Viglino P, et al. (1996) A molecular code dictates sequence-specific DNA recognition by homeodomains. *EMBO J* 15: 4992–5000.
 93. Thisse C, Thisse B, Schilling TF, Postlethwait JH (1993) Structure of the zebrafish *snail1* gene and its expression in wild-type, spadetail and no tail mutant embryos. *Development* 119: 1203–1215.
 94. Nasevicius A, Ekker SC (2000) Effective targeted gene ‘knockdown’ in zebrafish. *Nat Genet* 26: 216–220.